JOURNAL OF ALGEBRA 2, 42-47 (1965)

Remarks on a Paper of Taussky*

RAYMOND REDHEFFER

Department of Mathematics, University of California, Los Angeles, California

Communicated by Marshall Hall, Jr.

Received June 17, 1964

1. OBJECTIVES

In this introductory section a, b, c, x, α, γ denote square matrices of the same order, with complex elements. The appropriate zero and unit matrices are 0 and 1 respectively. We write $a > b$ if $a - b$ is positive definite, $a \geqslant b$ if $a - b$ is positive semidefinite. With this understanding. Theorems 5 and 6 of [*1*] are together equivalent to :

THEOREM 1 (Taussky). *Let $a^n \to 0$ and $c > 0$. Then the equation $x - axa^* = c$ has a solution $x > 0$. If $c = 1$ every solution x, with $x > 0$, has all characteristic values $\geqslant 1$.*

The statement sharpens a theorem of Stein (see [1]) according to which the inequality $x - axa^* > 0$ has a solution $x > 0$, provided $a^n \to 0$.

A matrix α is said to be *stable* if all its characteristic values have negative real parts. Theorem 3 of [*1*] is:

THEOREM 2 (Taussky). *If α is stable and $\gamma > 0$, the equation*

$$\alpha x + x\alpha^* + \gamma = 0$$

has a solution $x > 0$.

This generalizes an important theorem of Lyapunov, viz., the corresponding statement with $\gamma = 1$. It is shown in [*1*] that Theorem 2 follows from Theorem 1, and that both follow from Lyapunov's theorem. However, an independent proof is not given.

As our first objective we give a proof of Theorems 1 and 2 that is wholly self-contained, and yet seems astonishingly elementary. The actual result

* The preparation of this material was sponsored by the Office of Naval Research. Reproduction in whole or in part is permitted for any purpose of the United States Government.

is a good deal sharper in several respects; and this sharpening is our second objective. It turns out that the proof applies not only to the algebra of matrices but to a broad class of significant algebraic structures; e.g., operators on Hilbert space, Banach algebras, and normed rings. This generalization of the underlying algebra is our third objective. Finally, we generalize the original equation.

2. The Algebra

Instead of matrices, let $a, b, c, x, y, a_i, b_i, \alpha, \beta, \gamma, x_i$ denote elements of an arbitrary ring, K, with a valuation $|\ |$. The latter is to be a nonnegative real-valued function defined on K and satisfying

$$|x+y| \leqslant |x| + |y|, \qquad |xy| \leqslant |x|\,|y|, \qquad |x| = 0 \Rightarrow x = 0$$

where the last "0" denotes the zero element in K. A statement such as $a^n \to 0$ means $|a^n| \to 0$, and similarly for all other statements which, directly or indirectly, involve limits. We assume K is complete; that is, Cauchy sequences converge. However, we do not require a unit, 1, or multiplication of elements of K by scalars. The additional algebraic structure introduced in Section 4 is used only to show the relation of our results to those of Lyapunov and Taussky.

3. A Simple Equation

We now establish:

Remark 1. *If $\Sigma\, |a^n|\,|b^n| < \infty$ the equation $x - axb = c$ has a unique solution, and the solution is $x = c + acb + a^2cb^2 + \cdots$.*
For proof, define a sequence by $x_n = c + ax_{n-1}b$, starting with x_0. Since

$$|x_{n+1} - x_n| \leqslant |a^n|\,|b^n|\,|ax_0b + c - x_0|,$$

$\Sigma\,(x_{n+1} - x_n)$ converges. This shows that $\lim x_n = x$ exists as $n \to \infty$, and thus produces a solution in the desired explicit form. If $\tilde{x}$ is another solution the choice $x_0 = \tilde{x}$ gives a sequence of the foregoing type. On the one hand $x_n = \tilde{x}$, since $\tilde{x} = c + a\tilde{x}b$, and on the other hand $x_n \to x$ by the above. Therefore $\tilde{x} = x$, establishing uniqueness.

Remark 2. *Suppose there are integers p and q such that $|a^p| \leqslant 1$ and $|b^q| < 1$, or such that $|a^p|\,|b^p| < 1$. Then the conclusion of Remark 1 holds.*
Indeed, if $|a^p| \leqslant 1$, the equation $|a^n| = |a^{pk+r}| \leqslant |a^r|$ shows that

$| a^n |$ is bounded for all n, and similarly if $| b^q | < 1$, then $| b^n | \leqslant M\theta^n$ for some constants M and $\theta < 1$. Thus $| a^n | \, | b^n | \leqslant \tilde{M}\theta^n$ for another constant $\tilde{M}$. The same conclusion follows (in the same way) if $| a^p | \, | b^p | < 1$, and the series in Remark 1 therefore converges.

4. The Theorems of Taussky and Lyapunov

We now introduce a function $*$, on K into K, such that

$$(ab)^* = b^*a^*, \qquad (a + b)^* = a^* + b^*, \qquad | a^* | \leqslant | a | .$$

(The function need not be an involution and hence there is always the choice $a^* \equiv 0$). We say $a \geqslant b$ if $a - b$ has the form pp^*, or is a sum of terms of this form, or a limit of such sums. Since $s = \lim s_n$ implies $asa^* = \lim as_na^*$, and since $app^*a^* = (ap)(ap)^*$, we get the following result, by inspection of the series in the conclusion of Remark 1:

Remark 3. *Let* $\lim a^n = 0$. *Then the equation* $x - axa^* = c$ *has a unique solution; and* $c \geqslant 0 \Rightarrow x \geqslant c$.

If K has a unit, the choice $c = 1$ gives $x \geqslant 1 + aa^* + \cdots$. Thus, Remark 3 sharpens both parts of Theorem 1.

Technically speaking, the unit need not be two-sided or $\geqslant 0$ in the foregoing. We now assume it is two-sided, and state:

Remark 4. *Let* $a = (1 - \alpha)^{-1}(1 + \alpha)$ *and* $b = (1 + \beta)(1 - \beta)^{-1}$ *satisfy the hypothesis of Remark 2. Then:*

(i) *If* $(1 - \alpha)^{-1}$ *is a left inverse and* $(1 - \beta)^{-1}$ *a right inverse, the equation* $2(\alpha x + x\beta) + \gamma = 0$ *has at most one solution.*

(ii) *If* $(1 - \alpha)^{-1}$ *is a right inverse and* $(1 - \beta)^{-1}$ *a left inverse, the equation* $2(\alpha x + x\beta) + \gamma = 0$ *has at least one solution; namely* $x = c + acb + a^2cb^2 + \cdots$, *where* $c = (1 - \alpha)^{-1}\gamma(1 - \beta)^{-1}$.

In view of the uniqueness and existence asserted in Remark 2, the result follows by inspection of the equation

$$(1 - \alpha)\, x(1 - \beta) - (1 + \alpha)\, x(1 + \beta) = \gamma,$$

as in [*1*]. Naturally, the function $*$ is not needed in Remark 4.

Continuing with our assumption that 1 is two-sided, we now introduce the function $*$ again, and assume $1^* = 1$. Thus, an inverse for x gives one for x^*. The element α is said to be *stable* if $1 - \alpha$ has a two-sided inverse and

$$| (1 - \alpha)^{-1}(1 + \alpha) | < 1.$$

It is easily shown that this holds for matrices stable in the sense of Lyapunov, and hence the following remark contains Theorem 2:

REMARK 5. *If α is stable, the equation* $2(\alpha x + x\alpha^*) + \gamma = 0$ *has a unique solution; and* $\gamma \geqslant 0 \Rightarrow x \geqslant (1 - \alpha)^{-1} \gamma (1 - \alpha^*)^{-1}$.

5. More General Equations

Let f be a continuous function on K into K and define a sequence by $x_n = f(x_{n-1})$, starting with $x_0 \in K$. If the condition

$$\lim x_n = \lim x_{n-1} = x \qquad (n = n_i \to \infty)$$

holds for a particular sequence $\{n_i\}$ then x is a fixed point, $x = f(x)$. Conversely, every fixed point is given by a sequence of that sort. This trivial observation establishes the equivalence of (i) and (ii) in the following remark. The equivalence with (iii) follows from the fact (pointed out by Ernst Straus) that $x - Tx = c$ has at most one solution, when $\lim T^n x = 0$:

REMARK 6. *Let T be a continuous linear transformation on K into K such that* $\lim T^n x = 0$ *for each $x \in K$. Then the following statements are equivalent:*

(i) *x is a solution of $x - Tx = c$,*

(ii) *x is a limit point of the sequence $\{c + Tc + T^2c + \cdots + T^n c\}$,*

(iii) $x = c + Tc + T^2c + \cdots$.

Remark 6 really requires only the additive structure on K, but the full structure is needed in the following.

We now consider the transformation T defined by

$$Tx = a_1 x b_1 + a_2 x b_2 + \cdots + a_m x b_m \,. \tag{1}$$

An easy use of mathematical induction shows that the iterated transformation is described by

$$T^n x = \Sigma\, (a_i a_j \cdots a_p a_q)\, x (b_q b_p \cdots b_j b_i).$$

The product of a's has n factors, the corresponding product of b's is obtained from this by writing b instead of a and reversing the order, and the sum is over all such products. That is, the sum is over all words of length n in the a's, repeated letters being allowed.

We shall use the symbol A^n to denote a word of n letters in the a's, and B^n for a word of n letters in the b's. When A^n and B^n occur in the same term we consider B^n to be a function of A^n, given by the process just described.

For example, if $A^4 = (a_2 a_7 a_3 a_3)$, then

$$A^4 x B^4 = (a_2 a_7 a_3 a_3) x (b_3 b_3 b_7 b_2).$$

In this notation the foregoing equation becomes

$$T^n x = \Sigma\, A^n x B^n \tag{2}$$

where the sum is over all words of length n in the a's. By inspection of $| T^n x |$ in (2), applying Remark 6, we get:

REMARK 7. *With T as in (1) let $\Sigma\Sigma \mid A^n \mid \mid B^n \mid < \infty$, where the inner sum is over all words of length n in the a's and the outer sum is over n. Then the equation $x - Tx = c$ has a unique solution, and the solution is*

$$x = c + Tc + T^2 c + \cdots .$$

6. Further Discussion of the General Case

The equation $A^{m+n} = A^m A^n$ is a brief way of saying that every word of length $m + n$ is a product of words of lengths m and n, and conversely. If $| A^p | \leqslant 1$ for a particular integer p and for all words A^p we easily show that $| A^n |$ is bounded. Similarly, if $| B^q | \leqslant \theta^q$ for a particular integer q and all words B^q we have

$$| B^n | \leqslant | B^q |^k \, | B^j | \leqslant \theta^{kq} \, | B^j | \leqslant M_0 \theta^n$$

for some constant M_0 . Since there are m^n terms in the sum (2), we deduce

$$| T^n x | \leqslant | x | \, M (m\theta)^n$$

for some constant M; and $m\theta < 1$ ensures convergence of the series in Remark 7.

Suppose next that $m^p \, | A^p | \, | B^p | < 1$ for some p and all words A^p. In this case (2) gives $| T^p x | \leqslant | x | \, \theta^p$, with $\theta < 1$. Since

$$T^n x = T^{kp+j} x = (T^p)^k \, (T^j x)$$

we get $| T^n x | \leqslant M \theta^n \, | x |$ for some constant M. Thus we have the following generalization of Remark 2:

REMARK 8. *With T as in (1), suppose there are integers p and q such that $| A^p | \leqslant 1$ for all words A^p, and $| B^q | < m^{-q}$ for all words B^q. Or else suppose $| A^p | \, | B^p | < m^{-p}$ for all words A^p. Then the conclusion of Remark 7 holds.*

In the notation of Section 4 we set $b_i = a_i{}^*$ and get the following generalization of Theorem 1:

REMARK 9. *Suppose there is an integer p such that all words A^p satisfy $| A^p | < m^{-p/2}$. Then the equation*

$$x - (a_1 x a_1{}^* + a_2 x a_2{}^* + \cdots + a_m x a_m{}^*) = c$$

has a unique solution; and $c \geqslant 0 \Rightarrow x \geqslant c$.

REFERENCE

1. TAUSSKY, OLGA. Matrices C with $C^n \to 0$. *J. Algebra* 1 (1964), 5-10. Other references are given here.